JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **163**, 541–558 (1992)

Weak Convergence and the Prokhorov Radius

GEORGE A. ANASTASSIOU

Department of Mathematical Sciences, Memphis State University, Memphis, Tennessee 38152

Submitted by Dorothy Maharam Stone

Received March 16, 1990

The *Prokhorov radius* for a set of probability measures satisfying basic moment conditions is introduced, through the Prokhorov distance of these measures from the Dirac measure at a fixed point of the real line. This is calculated precisely by the use of standard tools from the Kemperman geometric moment theory. The above radius gives the exact rate of weak convergence of these measures to the Dirac measure.

1. INTRODUCTION

Here we consider probability measures μ on $\mathbf{R}$ such that both $\int |t|\, d\mu$, $\int t^2\, d\mu < \infty$. We consider

$$M(\varepsilon_1, \varepsilon_2) = \left\{\mu : \left|\int t^j\, d\mu - \alpha^j\right| \leqslant \varepsilon_j, j = 1, 2\right\}, \tag{1}$$

where α is a given point in $\mathbf{R}$, also $0 < \varepsilon_j < 1$, $j = 1, 2$, and $0 < \varepsilon_2 + 2|\alpha|\, \varepsilon_1 < 1$. We would like to measure the "size" of $M(\varepsilon_1, \varepsilon_2)$ to be given by a simple formula involving only $\varepsilon_1, \varepsilon_2, \alpha$.

Since weak convergence of probability measures is of central importance and their standard weak topology is well described by the Prokhorov distance π, it is natural to define the *Prokhorov radius* for $M(\varepsilon_1, \varepsilon_2)$ as

$$D = \sup_{\mu \in M(\varepsilon_1, \varepsilon_2)} \pi(\mu, \delta_\alpha), \tag{2}$$

where δ_α is the Dirac measure at α. Using a geometric moment theoretical method due to Kemperman [2] we are able to calculate the exact value of D.

It will be helpful to mention.

0022-247X/92 $3.00

DEFINITION 1 (see [3]). Let U be a Polish space with a metric d and C be the set of all nonempty closed subsets of U. Let $A \in C$; then for $\varepsilon > 0$

$$A^{\varepsilon} = \{x : d(x, A) < \varepsilon\}.$$

Consider μ, ν probability measures on U. Prokhorov (1956) introduced his famous metric

$$\pi(\mu, \nu) = \inf\{\varepsilon > 0 : \mu(A) \leqslant \nu(A^{\varepsilon}) + \varepsilon,\ \nu(A) \leqslant \mu(A^{\varepsilon}) + \varepsilon,\ \forall A \in C\}.$$

When μ is a probability measure on $\mathbf{R}$, then

$$\pi(\mu, \delta_{\alpha}) = \inf\{r > 0 : \mu([\alpha - r,\ \alpha + r]) \geqslant 1 - r\}, \qquad \alpha \in \mathbf{R}. \tag{3}$$

Remark 1. One can restate that

$$D = \sup_{\mu \in M(\varepsilon_1, \varepsilon_2)} (\inf\{r > 0 : \mu([\alpha - r, \alpha + r]) \geqslant 1 - r\}). \tag{4}$$

Thus

$$r \geqslant D \qquad \text{iff } \mu([\alpha - r, \alpha + r]) \geqslant 1 - r, \text{ all } \mu \in M(\varepsilon_1, \varepsilon_2) \tag{5}$$

iff

$$\lambda_r = \inf_{\mu \in M(\varepsilon_1, \varepsilon_2)} \mu([\alpha - r, \alpha + r]) \geqslant 1 - r.$$

Obviously $0 < D \leqslant 1$; therefore, we are interested only in $r \in (0, 1]$. One can easily see that

$$D = \min\{r \in (0, 1] : \inf_{\mu \in M(\varepsilon_1, \varepsilon_2)} \mu([\alpha - r, \alpha + r]) \geqslant 1 - r\}. \tag{6}$$

Remark 2. (i) When $\varepsilon_1, \varepsilon_2 \to 0$, then $\int t\, d\mu \to \alpha$ and $\int t^2\, d\mu \to \alpha^2$. Thus $\int (t - \alpha)^2\, d\mu \to 0$, implying that $\int |t - \alpha|\, d\mu \to 0$. Hence for arbitrarily small $\varepsilon > 0$ we have

$$\varepsilon\mu(\{t : |t - \alpha| > \varepsilon\}) \leqslant \int_{\{t : |t - \alpha| > \varepsilon\}} |t - \alpha|\, d\mu \leqslant \int |t - \alpha|\, d\mu \leqslant \varepsilon^2.$$

That is, $\mu(\{t : |t - \alpha| > \varepsilon\}) \leqslant \varepsilon$. Therefore $\pi(\mu, \delta_{\alpha}) \leqslant \varepsilon$ and $D = D(\varepsilon_1, \varepsilon_2, \alpha) \leqslant \varepsilon$. Clearly $\pi(\mu, \delta_{\alpha}) \to 0$ and $D \to 0$ as $\varepsilon_1, \varepsilon_2 \to 0$, giving us that $\mu \rightrightarrows \delta_{\alpha}$ weakly. The knowledge of D gives the rate of weak convergence of μ to δ_{α}.

(ii) Let $|\alpha| < 1$. If $D \geqslant |\alpha| > 0$, then we cannot have $D \to 0$ as $\varepsilon_1, \varepsilon_2 \to 0$. Thus we are interested in $D < |\alpha|$; in that case there are r such that $D \leqslant r < |\alpha|$. More precisely: If $0 < \alpha < 1$, then $D \leqslant r < \alpha$, giving us

$\alpha - r < \alpha$ and $(\alpha - r)^2 < \alpha^2$, always true for sufficiently small ε_1, ε_2. If $-1 < \alpha < 0$, then $D \leqslant r < -\alpha$, giving us $\alpha < \alpha + r$ and $(\alpha + r)^2 < \alpha^2$, always true for sufficiently small $\varepsilon_1, \varepsilon_2$.

(iii) If $|\alpha| \geqslant 1$, then $D \leqslant r \leqslant |\alpha|$, true for any $0 < \varepsilon_j < 1, j = 1, 2$. More precisely: If $\alpha \geqslant 1$, then $D \leqslant r \leqslant \alpha$, giving us $\alpha - r < \alpha$ and $(\alpha - r)^2 < \alpha^2$. If $\alpha \leqslant -1$, then $D \leqslant r \leqslant -\alpha$, giving us $\alpha < \alpha + r$ and $(\alpha + r)^2 < \alpha^2$.

Comments (ii) and (iii) are used in the proof of Theorem 1.

2. Results

Our main result has as follows:

THEOREM 1. (i) *If* $\alpha = 0$ *we get* $D = \varepsilon_2^{1/3}$.

(ii) *If* $|\alpha| \geqslant 1$ *we get* $D = (\varepsilon_2 + 2|\alpha|\varepsilon_1)^{1/3}$.

(iii) *For* $0 < |\alpha| < 1$ *and sufficiently small* $\varepsilon_1, \varepsilon_2$ *we obtain again*

$$D = (\varepsilon_2 + 2|\alpha|\varepsilon_1)^{1/3}.$$

Comment 1. Our basic tools for the proof of Theorem 1 will come from [2, namely see pp. 93–96, 101–104, 107, 111, 120, 121].

Proof of Theorem 1(i). Here we have: (ℓ) line through $(\overline{BA})$, $Y^* = (y_1^*, y_2^*)$, $Y^{**} = (0, y_2^*)$, area(I) $= \operatorname{conv}(\widehat{BC}, \overline{BC})$, area(II) $= \operatorname{conv}(B \overset{\triangle}{C} A)$, area(III) $= \operatorname{conv}(\widehat{AC}, \overline{AC})$, $g(x) = (x, x^2)$, $V = \operatorname{convg}(\mathbf{R})$, V^0 interior of V, and

$$\mathscr{R} = \{(y_1, y_2) : |y_1 - \alpha| \leqslant \varepsilon_1, |y_2 - \alpha^2| \leqslant \varepsilon_2\}.$$

Let $\varepsilon_2 > r^2$; then the moment point $(0, \varepsilon_2) \in \mathscr{R} \cap V^0 \subset M(\varepsilon_1, \varepsilon_2)$ is above the line (ℓ) in Fig. 1.

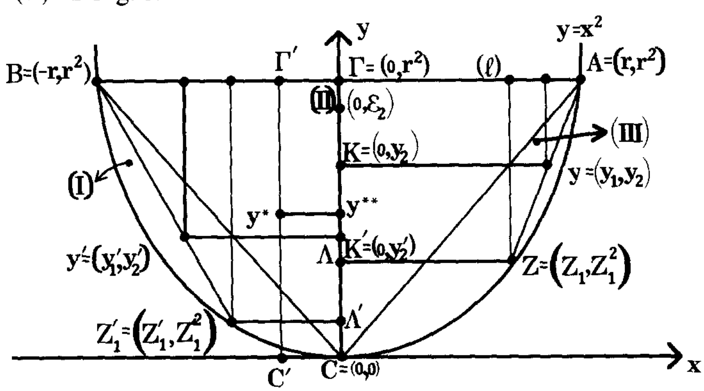

FIG. 1. $\alpha = 0$.

Hence

$$L := L_{[\alpha-r,\alpha+r]}(0,\varepsilon_2) := \inf_{\mu}(\mu([\alpha-r,\alpha+r]))=0,$$

where μ is such that

$$\int t\,d\mu=0 \qquad \text{and} \qquad \int t^2\,d\mu=\varepsilon_2$$

(see [2, pp. 96, 101, 120]). Therefore

$$\lambda_r=\inf_{\mu\in M(\varepsilon_1,\varepsilon_2)}\mu([\alpha-r,\alpha+r])=0\geqslant 1-r,$$

giving us $r\geqslant 1$ and $D=r=1$. Consequently, $\varepsilon_2>1$, which is a contradiction by the assumption that $\varepsilon_2<1$.

We have proved that $\varepsilon_2<r^2$ for all r such that $\lambda_r\geqslant 1-r$, that is $(0,\varepsilon_2)$ is below the line (ℓ). Furthermore any other moment point $(0,y_2)$ with $0<y_2\leqslant\varepsilon_2$ belongs to $\mathscr{R}\cap V^0$ and it is below the line (ℓ).

From [2, pp. 111, 120, 121]

$$L := L_{[\alpha-r,\alpha+r]}(y_1,y_2) := \inf_{\mu}(\mu([\alpha-r,\alpha+r])),$$

where μ is such that $\int t\,d\mu=y_1$, $\int t^2\,d\mu=y_2$ is given by the ratios

$$L=\frac{YA}{ZA}=\frac{K\Gamma}{\Lambda\Gamma}>\frac{K\Gamma}{C\Gamma},$$

where $Y=(y_1,y_2)\in\text{area(III)}$, and

$$L=\frac{Y'B}{Z'B}=\frac{K'\Gamma}{\Lambda'\Gamma}>\frac{K'\Gamma}{C\Gamma},$$

where $Y'=(y_1',y_2')$ belongs to area (I) (see Fig. 1).

Thus areas (I), (III) in terms of ratios are transferred into area (II). In fact as we can see from the above inequalities, in area (II) we get smaller L's. Therefore λ_r can be found from area (II).

Also observe that for any $Y^*\in\operatorname{conv}(BCA)$ we have

$$\frac{Y^*\Gamma'}{C'\Gamma'}=\frac{Y^{**}\Gamma}{C\Gamma}$$

(again see Fig. 1). So all that matters is the segment $C\Gamma$. Hence a typical $L=L(0,y_2)$ is given by the formula

$$L=\frac{K\Gamma}{C\Gamma}=\frac{r^2-y_2}{r^2}, \qquad \text{all} \quad 0<y_2\leqslant\varepsilon_2.$$

We would like to find the minimal $r \in (0, 1]$ such that $\lambda_r \geqslant 1 - r$. That is, we would like $L \geqslant 1 - r$, all $0 < y_2 \leqslant \varepsilon_2$

$$\text{iff } 1 - \frac{y_2}{r^2} \geqslant 1 - r, \qquad \text{all} \quad 0 < y_2 \leqslant \varepsilon_2$$

$$\text{iff } y_2 \leqslant r^3, \qquad \text{all} \quad 0 < y_2 \leqslant \varepsilon_2$$

$$\text{iff } \varepsilon_2 \leqslant r^3, \qquad \text{all} \quad \varepsilon_2^{1/3} \leqslant r.$$

Because the boundary points of $g(\mathbf{R})$ of the arc $B\widehat{C}A$ (see Fig. 1) play no role towards the calculation of D we get that

$$D = \min r = \varepsilon_2^{1/3}.$$

For the exclusion of the above boundary points see [2, pp. 102–104; Sect. 4 of noninterior points, also see the proofs of parts (ii) and (iii) of this theorem. The reasoning is exactly the same. ∎

To prove parts (ii) and (iii) of Theorem 1 we need

THEOREM 2. *Let μ be probability measures on $\mathbf{R}$ such that $\int t\, d\mu = y_1$, $\int t^2\, d\mu = y_2$. Let $\alpha \in \mathbf{R}$, $0 < r \leqslant 1$. Set*

$$L = L_{[\alpha - r, \alpha + r]}(y_1, y_2) = \inf_{\mu} (\mu([\alpha - r, \alpha + r])).$$

Then we find

(1) $L = 0$, *if* $y_2 + \alpha^2 - 2\alpha y_1 > r^2$;

(2) $L = 1$, *if* $y_2 = y_1^2$ *with* $y_1 \in [\alpha - r, \alpha + r]$.

(3) *Let* $y_2 \neq y_1^2$ *and* $r\,|y_1 - \alpha| \leqslant y_2 + \alpha^2 - 2\alpha y_1 \leqslant r^2$; *then*

$$L = 1 - \frac{[y_2 + \alpha^2 - 2\alpha y_1]}{r^2}.$$

(4) *Let* $y_2 \neq y_1^2$ *and* $y_2 + \alpha^2 - 2\alpha y_1 \leqslant r(y_1 - \alpha)$; *then*

$$L = \frac{((\alpha + r) - y_1)^2}{(\alpha + r)^2 - 2(\alpha + r)\, y_1 + y_2}.$$

(5) *Let* $y_2 \neq y_1^2$ *and* $y_2 + \alpha^2 - 2\alpha y_1 \leqslant r(\alpha - y_1)$; *then*

$$L = (y_1 - (\alpha - r))^2 / ((\alpha - r)^2 - 2(\alpha - r)\, y_1 + y_2).$$

Comment 2. Here $y_1^2 \leqslant y_2$ and $(y_1, y_2) \in \operatorname{conv} g(\mathbf{R})$, where $g(x) = (x, x^2)$, $x \in \mathbf{R}$.

Proof of Theorem 2.

Case (2). Set again $V = \text{convg}(\mathbf{R})$. We use the terminology and the results of [2, pp. 102–104; Sect. 4 of noninterior points]. Let y be a boundary point of V; then $V_y = \{y\}$. And $\Gamma^y := g(\mathbf{R}) \cap V_y = \{y\}$, i.e., $\Gamma^y = \{y\}$. Also $T^y := g^{-1}(\{y\}) = t_y$, with g being a one-to-one mapping. That is, $T^y = \{t_y\}$. From Theorem 9 of [2, p. 103] we have: Let y be a fixed boundary point of V and let μ be a probability measure on $\mathbf{R}$ such that $\int |t| \, d\mu < \infty$, $\int t^2 \, d\mu < \infty$ with $\int t \, d\mu = y_1$, $\int t^2 \, d\mu = y_2$; $y = (y_1, y_2)$. Then μ is concentrated at $\{t_y\}$. The conclusion is that $L = 1$.

Case (1). Let $y_1 \in S' := \mathbf{R} - [\alpha - r, \alpha + r]$; then $y = (y_1, y_1^2)$ is a boundary point of $g(S')$. Therefore any measure μ as in Case (2) will be concentrated at $\{y_1\}$. Consequently,

$$L = L_{[\alpha - r, \alpha + r]}(y_{\text{boundary}} \in g(S')) = 0.$$

Furthermore

$$L = L_{[\alpha - r, \alpha + r]}(y \in (\text{convg}(S'))^0) = 0.$$

By letting $W_{S'} = \overline{\text{convg}(S')}$, we have that

$$L = L_{[\alpha - r, \alpha + r]}(y) = 0, \qquad \forall y \in W_{S'}.$$

See also [2, pp. 101, 120, 121].

Case (3). Let $\alpha \geqslant 0$. When $\alpha < 0$, the proof is exactly the same; therefore, it is omitted (see Fig. 2).

Let $S = [\alpha - r, \alpha + r]$, $S' = \mathbf{R} - S$, $0 < r \leqslant 1$. Set area (II) $= \text{conv}\,(BCA)$, $W_S = \overline{\text{convg}\,(S)}$. Consider the line (ℓ) through $(\overline{BA})$, also call it (H'), and consider the line (λ) tangent to $y = x^2$ and parallel to (ℓ) that goes through the point $C = (\alpha, \alpha^2)$, also call it (H).

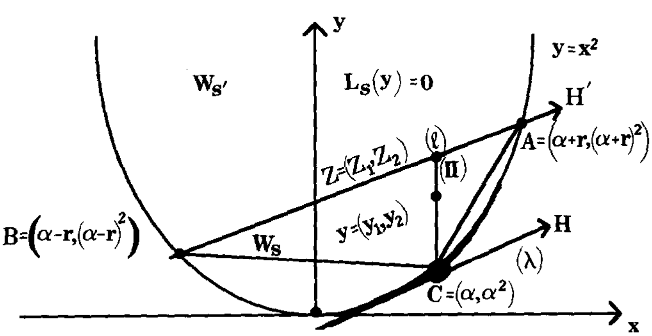

FIG. 2. The case where $y \in \text{conv}\,(BCA)$.

Here see [2, pp. 101, 102, 111, 120, 121]. One can easily prove that $y=(y_1, y_2)\in \operatorname{conv}(BCA)$ iff

$$r\mid y_1-\alpha\mid \leqslant y_2+\alpha^2-2\alpha y_1\leqslant r^2.$$

And line (ℓ) is given by

$$(-2\alpha)\,x+1y+(\alpha^2-r^2)=0,$$

i.e., $\tilde{A}=-2\alpha$, $\tilde{B}=1$, $\tilde{C}=\alpha^2-r^2$ for the line (ξ): $\tilde{A}x+\tilde{B}y+\tilde{C}=0$. We know that the distance

$$d((x_0, y_0), \xi)=\frac{|\tilde{A}x_0+\tilde{B}y_0+\tilde{C}|}{\sqrt{\tilde{A}^2+\tilde{B}^2}}.$$

Here

$$d((y_1, y_2), l)=\frac{|-2\alpha y_1+y_2+(\alpha^2-r^2)|}{\sqrt{4\alpha^2+1}}$$

and

$$d((\alpha, \alpha^2), \ell)=\frac{r^2}{\sqrt{4\alpha^2+1}}.$$

Therefore from [2]

$$L=L_S(y)=\frac{d((y_1, y_2), \ell)}{d((\alpha, \alpha^2), \ell)},$$

i.e.,

$$L=\frac{|-2\alpha y_1+y_2+(\alpha^2-r^2)|}{r^2}, \qquad \text{all } y=(y_1, y_2)\in \text{area(II)}.$$

Hence

$$L=L_S(y)=1-\frac{[y_2+\alpha^2-2\alpha y_1]}{r^2},$$

for all $y=(y_1, y_2)$ such that

$$r\mid y_1-\alpha\mid \leqslant y_2+\alpha^2-2\alpha y_1\leqslant r^2.$$

Cases (4, 5). Here we consider $\alpha\geqslant 0$; when $\alpha<0$, the proof is omitted as in Case (3) (see Fig. 3).

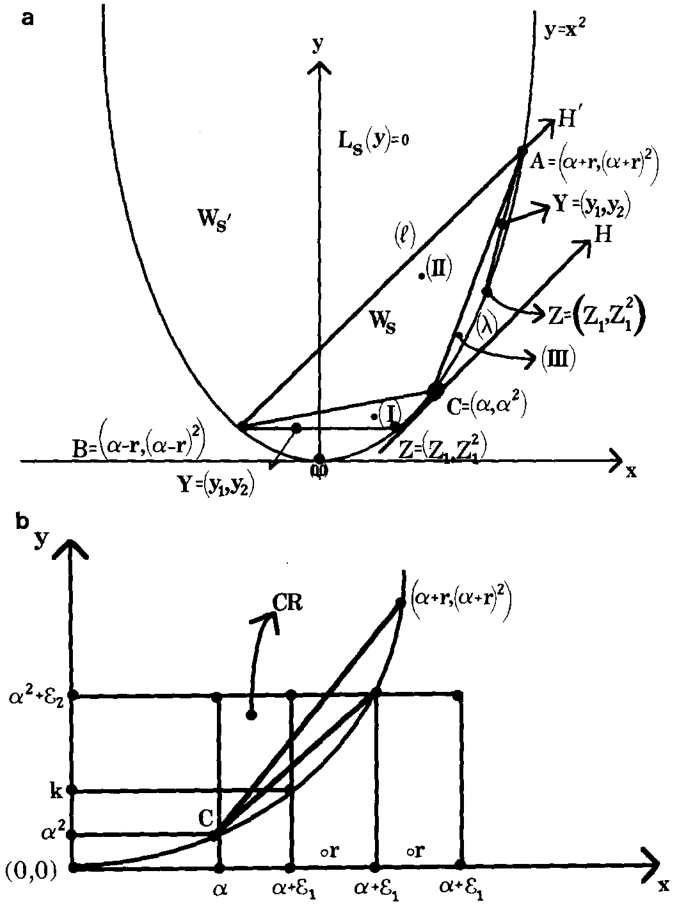

FIG. 3(a) The case where Y is below area (II). (b) The case where $L=1$.

Here $S=[\alpha-r,\alpha+r]$, $0<r\leqslant 1$, area (I) $=\operatorname{conv}(\widehat{BC},\overline{BC})$, and area (III) $=\operatorname{conv}(\widehat{AC},\overline{AC})$. Consider $W_S=\overline{\operatorname{convg}(S)}$, the line (ℓ) through $(\overline{BA})$, also call it (H'), and the line (λ) tangent to $y=x^2$ and parallel to (ℓ) that goes through the point $C=(\alpha,\alpha^2)$, also call it (H).

One can easily see that (y_1,y_2) is on or below (CA) iff $y_2+\alpha^2-2\alpha y_1\leqslant r(y_1-\alpha)$ $(y_1\geqslant\alpha)$. And (y_1,y_2) is on or below (BC) iff $y_2+\alpha^2-2\alpha y_1\leqslant r(\alpha-y_1)$ $(y_1\leqslant\alpha)$. Furthermore, (y_1,y_2) is on or above (BA) iff $y_2+\alpha^2-2\alpha y_1\geqslant r^2$. Again we use the techniques of [2, pp. 93–96, 101, 102, 111, 120, 121].

Area (III). From

$$L=L_S(y)=\frac{YA}{ZA}=\frac{\alpha+r-y_1}{\alpha+r-z_1}=\frac{(\alpha+r)^2-y_2}{(\alpha+r)^2-z_1^2}$$

we get

$$z_1 = \frac{y_1(\alpha+r) - y_2}{(\alpha+r) - y_1}$$

and

$$L = \frac{((\alpha+r) - y_1)^2}{(\alpha+r)^2 - 2(\alpha+r)\, y_1 + y_2}, \tag{7}$$

for all $Y = (y_1, y_2)$ such that

$$y_2 + \alpha^2 - 2\alpha y_1 \leqslant r(y_1 - \alpha) \qquad (y_1 \geqslant \alpha).$$

Area (I). The subcase $y_1 = \alpha - r = z_1$ is trivial. For $z_1 \neq -(\alpha - r)$ we get

$$L = L_S(y) = \frac{BY}{BZ} = \frac{y_1 - (\alpha - r)}{z_1 - (\alpha - r)} = \frac{y_2 - (\alpha - r)^2}{z_1^2 - (\alpha - r)^2}.$$

Thus

$$z_1 = \frac{y_2 - y_1(\alpha - r)}{y_1 - (\alpha - r)}$$

and

$$L = \frac{(y_1 - (\alpha - r))^2}{(\alpha - r)^2 - 2(\alpha - r)\, y_1 + y_2}, \tag{8}$$

for all $Y = (y_1, y_2)$ such that

$$y_2 + \alpha^2 - 2\alpha y_1 \leqslant r(\alpha - y_1) \qquad (y_1 \leqslant \alpha).$$

Remark 3. Assume $y_2 = (\alpha - r)^2$; then $z_1^2 = (\alpha - r)^2$ and $z_1 = \pm(\alpha - r)$. If $z_1 = \alpha - r$, then this is a trivial subcase. If $z_1 = -(\alpha - r)$, then

$$L = L_S(y) = \frac{y_1 - (\alpha - r)}{2(r - \alpha)}. \tag{9}$$

But (9) is covered by (8) for this particular case of $y_2 = (\alpha - r)^2$ and $z_1 = r - \alpha$.

Remark 4. Proving the case of $\alpha < 0$ we deal with the following: Assume $y_2 = (\alpha + r)^2$; then $z_1^2 = (\alpha + r)^2$ an $z_1 = \pm(\alpha + r)$. If $z_1 = \alpha + r$, then this is a trivial subcase. If $z_1 = -(\alpha + r)$, then

$$L = L_S(y) = \frac{(\alpha + r) - y_1}{2(\alpha + r)}. \tag{10}$$

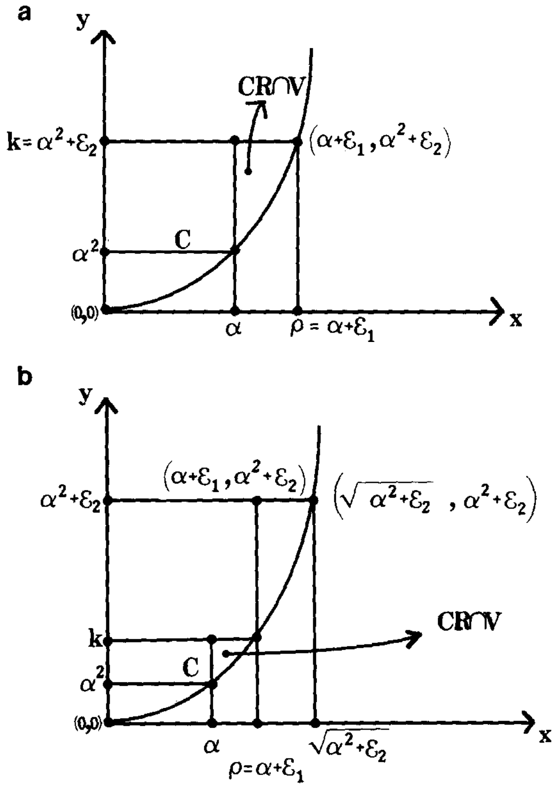

FIG. 4. (a) $\rho^2=k$. (b) $\alpha \leqslant \rho \leqslant \sqrt{\alpha^2+\varepsilon_2}$. (c) $\alpha \leqslant \rho \leqslant \alpha+\varepsilon_1$. (d) $\rho \leqslant \alpha+\varepsilon_1$.

Substituting $y_2=(\alpha+r)^2$, $z_1=-(\alpha+r)$ into (7) we get (10). That is, this subcase is covered by formula (7).

The above ends the proof of Theorem 2. ∎

The proof of Theorem 1 is continued.

Remark 5. (1) $L=1$ *means nothing towards the calculation of* D. Consider $\alpha \geqslant 0$; the case of $\alpha<0$ as similarly treated is omitted. We know from Theorem 2(2) that the points (y_1, y_2) such that $y_2=y_1^2$, $y_1 \in [\alpha-r, \alpha+r]$ produce $L=1$.

Consider the rectangle $\mathcal{R}$ of points (y_1, y_2) such that $|y_j-\alpha^j| \leqslant \varepsilon_j$, $j=1, 2$. That is, $\alpha-\varepsilon_1 \leqslant y_1 \leqslant \alpha+\varepsilon_1$ and $\alpha^2-\varepsilon_2 \leqslant y_2 \leqslant \alpha^2+\varepsilon_2$. Note that $\alpha+\varepsilon_1>\alpha$ and $\alpha^2+\varepsilon_2>\alpha^2$. Again let $V=\operatorname{convg}(\mathbf{R})$; $g(x)=(x, x^2)$. Obviously $\mathcal{R} \cap V \neq \varnothing$, containing some of the interior points of V, call them (y_1^0, y_2^0).

Note that $y_1^2 \leqslant y_2$ (see Fig. 4).

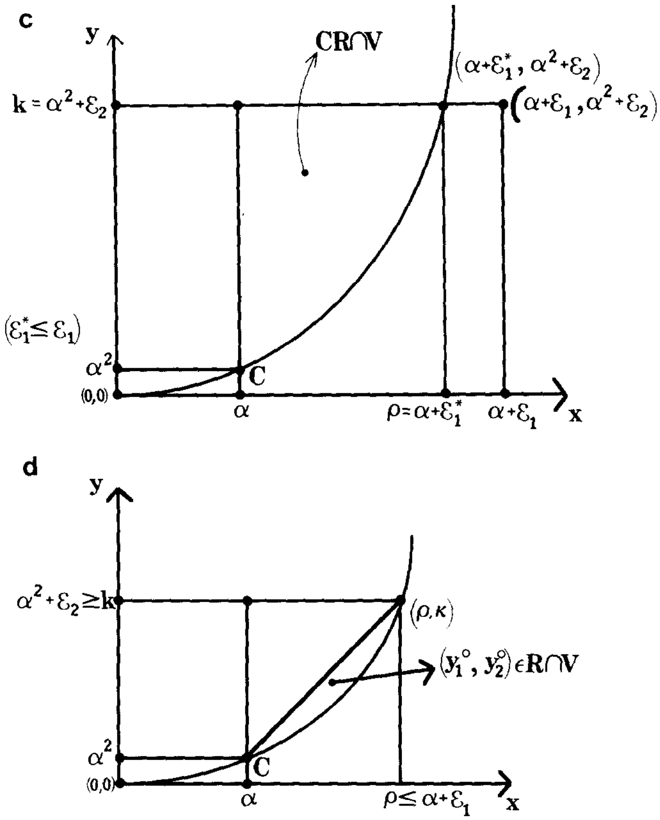

FIG. 4—*Continued*

The point $\rho=\alpha+\varepsilon_1^*$, $\varepsilon_1^* \leqslant \varepsilon_1$. Thus for *any* $r \in (0, 1]$ we have $\Gamma := [\alpha, \alpha+r] \cap [\alpha, \alpha+\varepsilon_1^*] \neq \varnothing$. If $r \leqslant \varepsilon_1^*$, $\Gamma=[\alpha, \alpha+r]$, and if $r \geqslant \varepsilon_1^*$, $\Gamma=[\alpha, \alpha+\varepsilon_1^*]$. If $r \leqslant \varepsilon_1^*$, then

$$\operatorname{convg}([\alpha, \alpha+r]) \subset \operatorname{convg}([\alpha, \alpha+\varepsilon_1^*]).$$

If $r \geqslant \varepsilon_1^*$, then

$$\operatorname{convg}([\alpha, \alpha+r]) \supset \operatorname{convg}([\alpha, \alpha+\varepsilon_1^*]).$$

In any case there exist interior points of V: $(y_1^0, y_2^0) \in \mathscr{R} \cap V$ which also belong to $\operatorname{convg}([\alpha, \alpha+r]) \cap \operatorname{convg}([\alpha, \alpha+\varepsilon_1^*]) \neq \varnothing$. Fix such (y_1^0, y_2^0). By Lemma 2 [2, p. 96] there exists a probability measure μ^0 on $\mathbf{R}$ such that

$$\int t\,d\mu^0 = y_1^0, \qquad \int t^2\,d\mu^0 = y_2^0. \tag{11}$$

Obviously

$$\operatorname{convg}([\alpha, \alpha+r]) \cap \operatorname{convg}([\alpha, \alpha+\varepsilon_1^*]) \subset \operatorname{convg}([\alpha, \alpha+r])$$
$$\subset \operatorname{convg}([\alpha-r, \alpha+r]).$$

That is, (y_1^0, y_2^0) is an interior point of convg $([\alpha-r, \alpha+r])$, for which we find from Theorem 2 that $\inf_\mu \mu([\alpha-r, \alpha+r]) < 1$, μ as in (11), for any $r \in (0, 1]$.

Therefore

$$\lambda_r = \inf_{\mu \in M(\varepsilon_1, \varepsilon_2)} \mu([\alpha-r, \alpha+r]) < 1 = L.$$

$L=1$ comes from the associated boundary points in that neighborhood. *Consequently, the boundary points* (y_1, y_1^2), $y_1 \in [\alpha-r, \alpha+r]$, *do not contribute anything towards the calculation of* $D \in (0, 1]$. As such, they are excluded from consideration.

(2) $L=0$ *means nothing towards the calculation of* D. Consider $\alpha \geqslant 0$; the case of $\alpha<0$ as similarly treated is omitted. Let $L=0$; this can happen iff $y_2+\alpha^2-2\alpha y_1 \geqslant r^2$, with $(y_1, y_2) \neq ((\alpha \pm r), (\alpha \pm r)^2)$. Here $\alpha-\varepsilon_1 \leqslant y_1 \leqslant \alpha+\varepsilon_1$, $\alpha^2-\varepsilon_2 \leqslant y_2 \leqslant \alpha^2+\varepsilon_2$, $r \in (0, 1]$. Therefore $\lambda_r = 0 \geqslant 1-r$ can only be true for $r=1$, that is, $D=1$.

Note that

$$y_2+\alpha^2 \leqslant 2\alpha^2+\varepsilon_2 \tag{12}$$

and

$$-2\alpha y_1 \leqslant 2\alpha\varepsilon_1 - 2\alpha^2. \tag{13}$$

Adding (12), (13) we obtain

$$1 = r^2 \leqslant y_2+\alpha^2-2\alpha y_1 \leqslant \varepsilon_2+2\alpha\varepsilon_1.$$

Hence $\varepsilon_2+2\alpha\varepsilon_1 \geqslant 1$, which is a contradiction since we assumed $\varepsilon_2+2\alpha\varepsilon_1<1$.

The last assumption is justified by the following: Let $\varepsilon_2+2\alpha\varepsilon_1 \geqslant 1$ and $\varepsilon = \max(\varepsilon_1, \varepsilon_2)$. Then $1 \leqslant \varepsilon_2+2\alpha\varepsilon_1 \leqslant \varepsilon(1+2\alpha)$, giving us $\varepsilon \geqslant (1+2\alpha)^{-1}$, that is ε is bounded away from zero. But we would like to have $\varepsilon_1, \varepsilon_2 \to 0$, that is, it should be $\varepsilon \to 0$. Consequently the case of $\varepsilon_2+2\alpha\varepsilon_1 \geqslant 1$ is absurd.

We have established that $L=0$ plays no role in the calculation of D. That is, we have proved that *the points* $(y_1, y_2) \in \mathscr{R} \cap V - \{((\alpha \pm r), (\alpha \pm r)^2)\}$, *which are on or above the line through* $\{((\alpha \pm r), (\alpha \pm r)^2)\}$, *play no role in the calculation of* D. As such, they are excluded from consideration.

The proof of Theorem 1 is continued.

Remark 6. The case where $\alpha > 0$: Note that any $(y_1, y_2) \in \mathscr{R} \cap V$ fulfills

$$y_1 \geqslant \max(\alpha - \varepsilon_1, -\sqrt{\alpha^2 + \varepsilon_2}).$$

Take $\bar{y}_1 = \max(\alpha - \varepsilon_1, -\sqrt{\alpha^2 + \varepsilon_2})$, $\bar{y}_2 = \alpha^2 + \varepsilon_2$. Obviously $(\bar{y}_1, \bar{y}_2) \in \mathscr{R} \cap V$ and there exists a probability measure μ on $\mathbf{R}$ such that

$$\int t \, d\mu = \bar{y}_1, \qquad \int t^2 \, d\mu = \bar{y}_2$$

(see [2, Lemma 2, p. 96]).

Consider any *admissible* $0 < r \leqslant 1$, i.e., $\lambda_r \geqslant 1 - r$. Assume $\bar{y}_2 + \alpha^2 - 2\alpha\bar{y}_1 > r^2$; then

$$2\alpha^2 + \varepsilon_2 - 2\alpha\bar{y}_1 > r^2.$$

If $\bar{y}_1 = \alpha - \varepsilon_1$, then $\varepsilon_2 + 2\alpha\varepsilon_1 > r^2$. If $\bar{y}_1 = -\sqrt{\alpha^2 + \varepsilon_2}$, then $(\alpha + \sqrt{\alpha^2 + \varepsilon_2})^2 > r^2$, then $r < \alpha + \sqrt{\alpha^2 + \varepsilon_2} \leqslant \alpha + \varepsilon_1 - \alpha$. Hence $\varepsilon_1 > r$.

Furthermore from Theorem 2 we obtain

$$\lambda_r = L = 0 \geqslant 1 - r; \qquad \text{then } r \geqslant 1.$$

Hence $r = 1$. Consequently we get either $\varepsilon_2 + 2\alpha\varepsilon_1 > 1$ or $\varepsilon_1 > 1$. Both contradict the assumption of the theorem.

Therefore $(\bar{y}_1, \bar{y}_2)$ *is on or below the line* (ℓ) *through the points* $\{((\alpha \pm r), (\alpha \pm r)^2)\}$. That is, $\bar{y}_2 + \alpha^2 - 2\alpha\bar{y}_1 \leqslant r^2$, for any admissible r.

Remark 6′. The case where $\alpha < 0$: Note that any $(y_1, y_2) \in \mathscr{R} \cap V$ fulfills

$$y_1 \leqslant \min(\alpha + \varepsilon_1, \sqrt{\alpha^2 + \varepsilon_2}).$$

Take $\bar{y}_1 = \min(\alpha + \varepsilon_1, \sqrt{\alpha^2 + \varepsilon_2})$, $\bar{y}_2 = \alpha^2 + \varepsilon_2$. Obviously $(\bar{y}_1, \bar{y}_2) \in \mathscr{R} \cap V$ and there exists a probability measure μ on $\mathbf{R}$ such that

$$\int t \, d\mu = \bar{y}_1, \qquad \int t^2 \, d\mu = \bar{y}_2$$

(see [2, Lemma 2, p. 96]).

Consider any *admissible* $0 < r \leqslant 1$, i.e., $\lambda_r \geqslant 1 - r$. Assume $\bar{y}_2 + \alpha^2 - 2\alpha\bar{y}_1 > r^2$; then $2\alpha^2 + \varepsilon_2 - 2\alpha\bar{y}_1 > r^2$. If $\bar{y}_1 = (\alpha + \varepsilon_1)$, then $\varepsilon_2 - 2\alpha\varepsilon_1 > r^2$. If $\bar{y}_1 = \sqrt{\alpha^2 + \varepsilon_2}$, then $(\alpha - \sqrt{\alpha^2 + \varepsilon_2})^2 > r^2$, then $\sqrt{\alpha^2 + \varepsilon_2} - \alpha > r$, then $\alpha + \varepsilon_1 \geqslant \sqrt{\alpha^2 + \varepsilon_2} > r + \alpha$. Hence $\varepsilon_1 > r$.

Furthermore from Theorem 2 we have $\lambda_r = L = 0 \geqslant 1 - r$; then $r \geqslant 1$; then

$r=1$. Thus, we get either $\varepsilon_2 - 2\alpha\varepsilon_1 > 1$ or $\varepsilon_1 > 1$. Both contradict the assumptions of the theorem.

Therefore $(\bar{y}_1, \bar{y}_2)$ *is on or below the line* (ℓ) *through the points* $\{((\alpha \pm r), (\alpha \pm r)^2)\}$. That is, $\bar{y}_2 + \alpha^2 - 2\alpha\bar{y}_1 \leqslant r^2$, for any admissible r.

The proof of Theorem 1 is continued.

Remark 7. The case where $\alpha > 0$:

(1) From Remark 2(ii) and (iii) we have: when $0 < \alpha < 1$ it holds that $(\alpha - r)^2 < \alpha^2$, for sufficiently small $\varepsilon_1, \varepsilon_2$; and when $\alpha \geqslant 1$ it holds again that $(\alpha - r)^2 < \alpha^2$, for any $0 < \varepsilon_j < 1, j = 1, 2$. Let

$$\bar{y}_1 = \max(\alpha - \varepsilon_1, -\sqrt{\alpha^2 + \varepsilon_2}), \qquad \bar{y}_2 = \alpha^2 + \varepsilon_2.$$

Then $(\bar{y}_1, \bar{y}_2)$ is above line (BC), in fact even strictly above line (ΦC) (see Fig. 5).

(2) Let $Y^* = (y_1^*, y_2^*) \in \mathscr{R} \cap \text{interior (area (I))} \cap V^0$. That is $|y_1^* - \alpha| \leqslant \varepsilon_1$, $|y_2^* - \alpha^2| \leqslant \varepsilon_2$. As before, from [2] we get that

$$L = \frac{BY^*}{BZ}.$$

Consider ZC and Y^{**} on BC such that Y^*Y^{**} is parallel to ZC. Then

$$\frac{BY^*}{BZ} = \frac{BY^{**}}{BC} \qquad \text{and} \qquad L = \frac{BY^{**}}{BC}.$$

Here $Y^{**} = (y_1^{**}, y_2^{**}) \in V^0$. One can easily see that $y_1^* < y_1^{**} < \alpha$, $y_2^* < y_2^{**} < \alpha^2$. Therefore $|y_1^{**} - \alpha| \leqslant \varepsilon_1$, $|y_2^{**} - \alpha^2| \leqslant \varepsilon_2$, i.e., $(y_1^{**}, y_2^{**}) \in \mathscr{R}$.

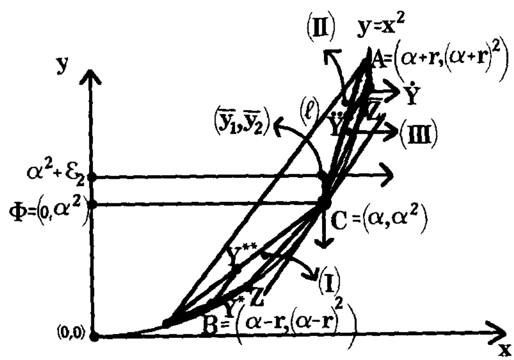

FIG. 5. $(\bar{y}_1, \bar{y}_2)$ is above line (ΦC).

Conclusion. Area (I) is transferred into area (II) and then all is taken care of by the formula of area (II) for L. See Theorem 2(3).

(3) Let $\dot{Y}=(\dot{y}_1, \dot{y}_2)\in \mathscr{R}\cap$ interior (area (III)) $\cap V^0$. That is, $|\dot{y}_1-\alpha| \leqslant \varepsilon_1$, $|\dot{y}_2-\alpha^2|\leqslant \varepsilon_2$. As before, from [2] we get that

$$L=\frac{A\dot{Y}}{A\bar{Z}}.$$

Consider $\bar{Z}C$ and $\ddot{Y}$ on AC such that $\ddot{Y}\dot{Y}$ is parallel to $C\bar{Z}$. Then

$$\frac{A\ddot{Y}}{AC}=\frac{A\dot{Y}}{A\bar{Z}} \qquad \text{and} \qquad L=\frac{A\ddot{Y}}{AC}.$$

Here $\ddot{Y}=(\ddot{y}_1, \ddot{y}_2)\in V^0$. One can easily see that $\alpha<\ddot{y}_1<\dot{y}_1$ and $\alpha^2<\ddot{y}_2<\dot{y}_2$. Therefore $|\ddot{y}_1-\alpha|\leqslant\varepsilon_1$ and $|\ddot{y}_2-\alpha^2|\leqslant\varepsilon_2$, i.e., $(\ddot{y}_1, \ddot{y}_2)\in\mathscr{R}$.

Conclusion. Area (III) is transferred into area (II) and then all is taken care of by the formula of area (II) of L. See Theorem 2(3).

Final Conclusion ($\alpha>0$). D can be calculated only from area (II) by the use of Theorem 2(3).

Remark 8. The case where $\alpha>0$: Calculate D (Theorem 1, cases (ii), (iii)). Here, all $y_1\geqslant\bar{y}_1=\max(\alpha-\varepsilon_1, -\sqrt{\alpha^2+\varepsilon_2})=\alpha-\varepsilon_1$ and all $y_2\leqslant\bar{y}_2=\alpha^2+\varepsilon_2$. Thus, all

$$\Delta := y_2+\alpha^2-2\alpha y_1\leqslant(\alpha^2+\varepsilon_2)+\alpha^2-2\alpha(\alpha-\varepsilon_1)=\varepsilon_2+2\alpha\varepsilon_1 =: K,$$

i.e., $1>K\geqslant\Delta$. Here we find λ_r over conv (BCA). We would like that any

$$L=1-\frac{[y_2+\alpha^2-2\alpha y_1]}{r^2}\geqslant 1-r$$

iff $\Delta=[y_2+\alpha^2-2\alpha y_1]\leqslant r^3$, for all $(y_1, y_2)\in(\text{conv}\,(BCA)\cap\mathscr{R}\cap V^0)$ iff $K\leqslant r^3$ iff $K^{1/3}\leqslant r$. Thus $D=\min r=K^{1/3}$.

Conclusion. $D=(\varepsilon_2+2\alpha\varepsilon_1)^{1/3}$ is true for $\alpha\geqslant 1$ with any $0<\varepsilon_j<1, j=1, 2$, and true for $0<\alpha<1$ when $\varepsilon_1, \varepsilon_2$ are sufficiently small.

The proof of Theorem 1 is continued.

Remark 7′. The case where $\alpha<0$:

(1) From Remark 2(ii), (iii) we have that when $-1<\alpha<0$ it holds that $(\alpha+r)^2<\alpha^2$, which is always true for sufficiently small $\varepsilon_1, \varepsilon_2$; and when $\alpha\leqslant -1$ it again holds that $(\alpha+r)^2<\alpha^2$, for any $0<\varepsilon_j<1$, $j=1, 2$. Let $\bar{y}_1=\min(\alpha+\varepsilon_1, \sqrt{\alpha^2+\varepsilon_2})$, $\bar{y}_2=\alpha^2+\varepsilon_2$. Then $(\bar{y}_1, \bar{y}_2)$ is above the line (AC) and, in fact, is even strictly above the line (ΦC) (see Fig. 6).

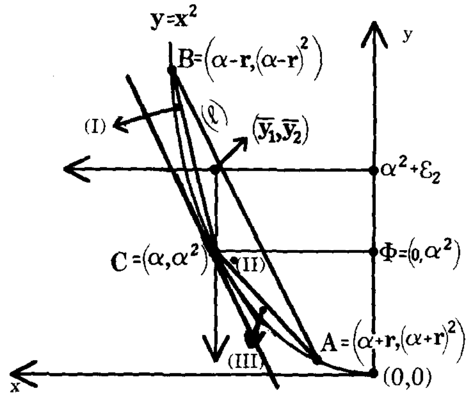

FIG. 6. $(\bar{y}_1, \bar{y}_2)$ is above line (AC).

(2) Let $Y^* = (y_1^*, y_2^*) \in \mathscr{R} \cap \text{interior (area (I))} \cap V^0$. That is, $|y_1^* - \alpha| \leqslant \varepsilon_1$, $|y_2^* - \alpha^2| \leqslant \varepsilon_2$. As before, from [2] we get that $L = BY^*/BZ$. Consider ZC and $Y^{**} = (y_1^{**}, y_2^{**})$ on BC such that $Y^* Y^{**}$ is parallel to ZC. Then

$$\frac{BY^*}{BZ} = \frac{BY^{**}}{BC}, \qquad \text{i.e., } L = \frac{BY^{**}}{BC}.$$

Here $Y^{**} \in V^0$ (see Fig. 7).

Also $y_1^* < y_1^{**} < \alpha$ and $\alpha^2 < y_2^{**} < y_2^*$, i.e., $|y_1^{**} - \alpha| \leqslant \varepsilon_1$ and $|y_2^{**} - \alpha^2| \leqslant \varepsilon_2$, giving us $(y_1^{**}, y_2^{**}) \in \mathscr{R}$.

Conclusion. Area (I) is transferred into area (II) and then all is taken care of by the formula of area (II) for L. See Theorem 2(3).

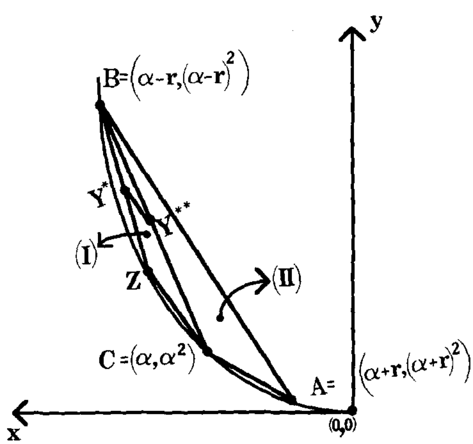

FIG. 7. Area (I) transferred into area (II).

(3) Let $\dot{Y}=(\dot{y}_1, \dot{y}_2)\in \mathscr{R}\cap$ interior (area (III)) $\cap V^0$. That is, $|\dot{y}_1-\alpha| \leqslant \varepsilon_1$ and $|\dot{y}_2-\alpha^2|\leqslant \varepsilon_2$. As before, from [2] we get that

$$L=\frac{A\dot{Y}}{A\bar{Z}}.$$

Consider $\bar{Z}C$ and $\ddot{Y}=(\ddot{y}_1, \ddot{y}_2)$ on AC such that $\ddot{Y}\dot{Y}$ is parallel to $C\bar{Z}$. Then

$$\frac{A\ddot{Y}}{AC}=\frac{A\dot{Y}}{A\bar{Z}}, \qquad \text{i.e., } L=\frac{A\ddot{Y}}{AC}.$$

Here $\ddot{Y}\in V^0$ (see Fig. 8).

Note that $\alpha<\ddot{y}_1<\dot{y}_1$ and $\dot{y}_2<\ddot{y}_2<\alpha^2$, that is, $|\ddot{y}_1-\alpha|\leqslant\varepsilon_1$ and $|\ddot{y}_2-\alpha^2|\leqslant\varepsilon_2$, i.e., $(\ddot{y}_1, \ddot{y}_2)\in\mathscr{R}$.

Conclusion. Area (III) is transferred into area (II) and then all is taken care of by the formula of area (II) for L. See Theorem 2(3).

Final Conclusion ($\alpha<0$). D can be calculated only from area (II), by the use of Theorem 2(3).

Remark 8′. The case where $\alpha<0$: Calculate D (Theorem 1, cases (ii), (iii)). Here, all $y_1\leqslant\bar{y}_1=\min(\alpha+\varepsilon_1, \sqrt{\alpha^2+\varepsilon_2})=\alpha+\varepsilon_1$ and all $y_2\leqslant\bar{y}_2=\alpha^2+\varepsilon_2$. Thus, all

$$\Delta := y_2+\alpha^2-2\alpha y_1\leqslant(\alpha^2+\varepsilon_2)+\alpha^2-2\alpha(\alpha+\varepsilon_1)=\varepsilon_2-2\alpha\varepsilon_1 =: K,$$

i.e., $1>K\geqslant\Delta$. Here we find λ_r over conv (BCA). We would like that any

$$L=1-\frac{[y_2+\alpha^2-2\alpha y_1]}{r^2}\geqslant 1-r$$

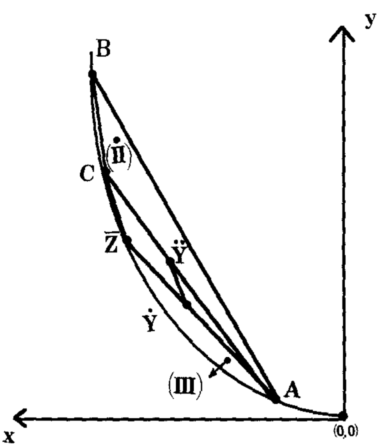

FIG. 8. Area (III) transferred into area (II).

iff $\Delta = [y_2 + \alpha^2 - 2\alpha y_1] \leqslant r^3$, for all $(y_1, y_2) \in (\text{conv}\,(BCA) \cap \mathscr{R} \cap V^0)$ iff $K \leqslant r^3$ iff $K^{1/3} \leqslant r$. Thus $D = \min r = K^{1/3}$.

Conclusion. $D = (\varepsilon_2 - 2\alpha\varepsilon_1)^{1/3}$ is true for $\alpha \leqslant -1$ with any $0 < \varepsilon_j < 1$, $j = 1, 2$, and true for $-1 < \alpha < 0$ when $\varepsilon_1, \varepsilon_2$ are sufficiently small. This completes the proof of Theorem 1. ∎

Acknowledgment

The author thanks Professor S. T. Rachev, of the University of California in Santa Barbara, for interesting discussions during the course of this work.

References

1. G. Anastassiou, The Levy radius of a set of probability measures satisfying basic moment conditions involving $\{t, t^2\}$, *Constr. Approx. J.* **3** (1987), 257–263.
2. J. H. B. Kemperman, The general moment problem, a geometric approach, *Annals of Mathematical Statistics* **39**, No. 1 (1968), 93–122.
3. S. T. Rachev and R. M. Shortt, Classification problem for probability metrics, *in* "Proceedings, Conference in Honor of Dorothy Maharam Stone, University of Rochester, Sept. 1987," A. M. S. Contemporary Mathematics, to appear.